Characterization of volatile compounds and microbial diversity of Arabica coffee in honey processing method based on different mucilage retention treatments

Faguang Hu [a,b,1], Haohao Yu [a,b,*], Xingfei Fu [a,b,1], Zhongxian Li [a], Wenjiang Dong [c], Guiping Li [a,b], Yanan Li [a,b], Yaqi Li [a,b], Bingqing Qu [a], Xiaofei Bi [a,b,*]

[a] *Institute of Tropical and Subtropical Cash Crops, Yunnan Academy of Agriculture Science, Baoshan 678000, Yunnan, China*
[b] *Yunnan Key Laboratory of Coffee, Baoshan 678000, China*
[c] *Spice and Beverage Research Institute, Chinese Academy of Tropical Agricultural Sciences, Wanning 571533, Hainan, China*

ARTICLE INFO

ABSTRACT

This study combined HS-SPME-GC–MS and high-throughput sequencing to explore how honey-processing methods with varying mucilage retentions impact volatile compounds and microbial communities in green coffee beans. HS-SPME-GC–MS revealed that the RH group (75 % to 80 % mucilage retention) had the highest relative content of volatile compounds. According to rOAV >1, 13 key aroma compounds were identified, contributing to flavors like "mellow" and "fruity". High-throughput sequencing identified seven dominant bacterial genera and four dominant fungal genera, with higher diversity of fungi than bacteria across treatments. Correlation analysis indicated that bacteria and fungi contribute to aroma formation, with bacteria more active in low-mucilage and fungi in high-mucilage treatments. Overall, the RH group was optimal for the aroma quality and bioactivity of green coffee beans. The findings of this research offers insights into aroma compound-microbe interactions in coffee mucilage fermentation, helping coffee producers optimize process parameters for better-quality coffee products.

1. Introduction

Coffee is renowned globally for its distinctive flavor and delightful aroma, ranking among the world's top three beverages alongside tea and cocoa. The three primary varieties of coffee include *Coffea arabica* (Arabica), *Coffea canephora* (Robusta), and *Coffea liberica* (Liberica). Among them, Arabica and Robusta are the most important commercial coffee varieties, accounting for about 70 % and 30 % of the world's agricultural coffee production, respectively (Bi et al., 2024). Arabica usually lives at high altitudes and is of good quality with a mild, delicate flavor. In contrast, Robusta can be grown in a wide range of areas, with a pronounced bitter and caramel flavor. In China, Yunnan province is the largest coffee-producing region, accounting for more than 98 % (almost all Arabica) of the country's cultivated area and production. Due to its favorable geographical location and natural conditions, Arabica produced here is known for its sweet aroma and aromatic acidity (Ma et al., 2022).

The sensory properties of coffee are the most important consumer parameter, described as a combination of aroma, flavor, texture, and mouthfeel (Seninde & Chambers, 2020). It's a complex process influenced by various factors such as genotype, cultivation method, post-harvest handling, roasting, and brewing quality (Pereira et al., 2019). Of these, one of the critical contributors to coffee beverage quality is the series of post-harvest handling performed to obtain dried beans suitable for roasting (Ferreira et al., 2023). These post-harvest handling involve peeling, degumming, fermentation, and drying, collectively called the coffee primary process. Wet processing is the most traditional method of coffee primary processing, where the mucilage is completely removed after peeling and then naturally dried, resulting in coffee beans with a bold flavor and bright acidity (Joët et al., 2010). Another traditional approach is dry processing, which involves using fresh coffee fruit exposed directly to the sun (Shofinita et al., 2024). Nowadays, as the coffee industry evolves, producers are exploring new primary processing methods to meet the market demand for coffees with different sensory

properties, aiming to increase their coffee quality and sensory differentiation. Most new primary processing methods are related to fermentation, such as honey treatments using mucilage for solar fermentation after a peeling process or adding an anaerobic fermentation process to wet processing or dry processing (Febrianto & Zhu, 2023). The application of distinct primary processing treatments gives rise to different metabolite characteristics. For instance, coffee with wet processing was characterized by chlorogenic acid, malic acid, and sucrose. Dry-processed coffee was represented with fructose and GABA, while honey-processed coffee possessed the intermediary metabolite profile between them and exhibited the highest antioxidant activity (Happyana et al., 2024). Similarly, honey processing has been shown to reduce the acrylamide content in roasted coffee compared to wet and dry processing, which may benefit human health (Vezzulli et al., 2022). For sensory properties, Aswathi et al. (2023) reported the metabolome of Robusta after honey processing; they reported that honey processing influenced carbohydrates, organic acids, and secondary metabolites, flavor precursors besides nonvolatile compounds that promote pleasant "sweet", "tea", "rose" and "chocolaty" flavors to the coffee. Zhai, Dong, Fu, et al., 2024, Zhai, Dong, Tang, et al., 2024) have shown that anaerobic dry processing increases amino acid and derivative content, as well as "fruity", "floral" and "sweet" aromas in roasted beans, meanwhile, honey processing increases lipid and phenolic acid content while enhancing "caramel", "roasted" aroma, and "butter" flavors. The formation of volatile compounds by mucilage during microbial fermentation is the underlying cause of this phenomenon, these compounds may contribute to the generation of additional flavors into the green beans and remain after the roasting process, including esters, alcohols, aldehydes, ketones, and terpenoids. Nevertheless, the lack of control over the fermentation process may result in a loss of quality control over the coffee product and even the potential formation of mycotoxins (Cardoso et al., 2021; Neto et al., 2018). Concerning this issue, it is crucial to understand the volatile kinetics that occur during mucilage fermentation. Consequently, quality control during the fermentation process should be further investigated.

Microbial activity is frequently accompanied by the fermentation process. Some microorganisms of coffee beans have proven to improve coffee flavor by producing aromatic flavor compounds through natural fermentation. In the wet process fermentation, yeast and lactic acid bacteria may significantly correlate to flavor compounds such as linalyl formate, linalool, cis-isoeugenol, trans-geraniol, and others (Todhanakasem et al., 2024). The microbial communities and produced metabolites change with the duration of the wet process fermentation, thus significantly affecting the biochemical composition of the coffee beans. Studies have shown that bacterial communities produce ester and alcohol by utilizing metabolic versatility, with *Enterobacter*, *Cluviraea*, and *Serratia* exhibiting strong connections with sugar and various volatile compounds, such as hexanal, benzaldehyde, 3-methylbenzaldehyde, 2-butenal, and 4-heptenal (Vale et al., 2024). Compared to wet processing, LAB and *Wallemia* were the dominant bacteria and fungi for honey processing, respectively, which may have given the beans more "sweet", "caramelly", "nutty", "pungent", and "hazelnut" flavors (Aswathi et al., 2022). Moreover, numerous studies have investigated the effects of exogenously introduced yeast on coffee. Lee et al. (2016) demonstrated that *R. oligosporus* fermentation of green coffee beans could induce modification of the aroma precursors of green coffee. Aswathi et al. (2024) suggested that the incorporation of *Saccharomyces cerevisiae* in honey-processed coffee results in the generation of "sweet", "fruity", and "caramel" flavors. Honey-processed coffee with anaerobic fermentation and yeast was characterized by aromas and flavors of "citrus", "caramel", "honey", "chocolate", and "chestnut" (Jimenez et al., 2023). Overall, microbial fermentation facilitates the production of specialty coffee with differentiated sensory properties.

As previously stated, numerous studies have examined the effects of different primary processing on flavor compounds in coffee beans. Conversely, there is a lack of research on quality control for individual primary processing methods. Among them, the role of microbial communities in coffee bean aroma formation is also indispensable. Accordingly, this study was conducted based on the assumption that different mucilage retention modifies the volatile compounds of green coffee beans through microbial fermentation, and these alterations can be explained as variations in flavor. In conclusion, this study aimed to determine the effects of different mucilage retention on microbial community composition and volatile compounds in honey processing using HS-SPME-GC–MS, and high-throughput sequencing. The relationship between microbial communities and volatile compounds was then explored. The objective of this study is to comprehensively understand quality control during Arabica coffee mucilage fermentation and to explore potential ways to convert the fermentation process into a more controlled and industrialized one.

2. Materials and methods

2.1. Sample preparation and collection

Coffee cherries (*Coffea arabica* cultivar "Catimor") were collected from Xinzhai Village, Lujiang Town, Baoshan City, Yunnan Province, China (25° 1′ 35″ N, 98° 49' 51″ E; altitude: 1200 m) in March 2023. The collected coffee cherries were carefully sorted to remove unripe, overripe, and defective cherries, to obtain cherries of consistent color and uniform size as test samples. The test samples were divided into five portions, each weighing 10 kg, for five treatments with different mucilage retention. A pre-test was conducted to determine the mucilage retention. That is, the coffee beans were weighed after peeling using a peeling machine (it is considered that retained mucilage is 100 % at this point), and the weight is recorded as *W1*. Then, the peeled coffee beans were weighed after being completely mechanically degummed using a degumming machine (it is considered that the mucilage is removed completely at this point), and the weight is recorded as *W2*. The equation is as follows (1):

$$W = (W1 - W2) \times R\% \tag{1}$$

In this context, *W* represents the quantity of single degumming, while *R*% denotes the rate of single degumming, which is regulated by adjusting the mechanical degumming machine loosening and tightening at a range of 20 % to 25 %. The pre-treatment of the formal test was consistent with the pre-test. Based on the results obtained from the pre-test, the quantity of degumming for the different treatments was reached according to the increase in the number of degumming times and verified by weighing. After degumming, let the beans stand for 40 min, and then manually stir with the sterile glove to distribute the mucilage evenly on each coffee-shelled bean. Subsequently, the coffee-shelled beans were exposed to sunlight to dry until the moisture content reached 11.0 ± 1.0 g/100 g DW, thereby preventing deterioration. The test samples were divided into five processing methods according to different mucilage retention: complete mucilage removal, control honey processing (CH); 20 % mucilage retention, white honey processing (WH); 50 % ∼ 60 % mucilage retention, yellow honey processing (YH); 75 % ∼ 80 % mucilage retention, red honey processing (RH); complete mucilage retention, black honey processing (BH), also a conventional honey processing method used in production. Finally, a hulling machine was used to remove the hulls and obtain raw coffee beans for further analysis.

2.2. HS-SPME-GC–MS analysis for volatile compounds

The green coffee bean samples were pulverized with liquid nitrogen and vortexed to ensure thorough mixing. 500 mg of each sample was weighed and mixed with saturated NaCl solution (Sinopharm Chemical Reagent Co., Ltd., Shanghai, China). Then, 20 μL (10 μg/mL) internal standard 3-hexanone-2,2,4,4-d4 solution (Sigma-Aldrich, St Louis, USA)

was added. The mixture was placed in headspace vials and the samples were extracted using a fully automated headspace solid phase micro-extraction (HS-SPME) technique. After shaking at 60 °C for 5 min, a 120 μm DVB/CWR/PDMS arrow (Agilent) was inserted into the sample headspace vial, and headspace extraction was performed for 15 min. The volatile compounds captured by the SPME fiber were injected into the GC apparatus (Model 8890, Agilent) at 250 °C for 5 min in the non-split mode, followed by separation and identification with gas chromatograph-mass spectrometer (GC–MS).

GC conditions were as follows: capillary column (DB-5MS, 30 m × 0.25 mm × 0.25 μm); carrier gas, high purity helium; constant flow rate, 1.2 mL/min; injector temperature, 250 °C; column temperature program, holding at 40 °C for 3.5 min, increasing to 100 °C at 10 °C/min, increasing to 180 °C at 7 °C/min, increasing to 280 °C at 25 °C/min, hold for 5 min. MS conditions are as follows: electron (energy, 70 eV) bombardment ion source (EI); EI temperature, 230 °C; quadrupole mass detector, 150 °C; transfer line temperature, 280 °C. The MS was selected ion monitoring (SIM) mode was used for the identification and quantification of analytes. The volatile compounds are separated step by step according to the melting and boiling points of the compounds, their polarity, and the force of interaction with the capillary column and other high-temperature resolution steps. When a sample molecule enters the ion source of a mass spectrometer, it undergoes ionization into ions due to the high energy action of the ion source. These ions are then separated in the mass analyzer according to m/z, resulting in the formation of a mass spectrogram. Volatile compounds were identified and quantified according to Metware's self-constructed high-coverage MS2 spectral tag library based on targeted spectra extraction algorithm, the mass spectrometry libraries of the National Institute of Standards and Technology (NIST17), multiple species, literature surveys, and several standards. The ion detection mode was selected for precision scanning, with one quantitative ion and two to three qualitative ions selected for each compound, respectively. The detected ions are identified according to the order of their peaks and within separate periods. The detected ions are identified according to the order of their peaks and within separate periods. If the retention time (RT) of the detected ions matched the standard reference, and the selected ions were present in the mass spectra of the samples after background subtraction, the substance was identified accordingly. Quantitative ions were then selected, and integration and correction processes were performed based on the peak area of these ions, which improved the accuracy of the quantification. Specifically refer to the method of Yuan et al. (2022). The relative content of the volatile compounds was calculated using eq. (2).

$$X_i = \frac{V_S \times C_s}{M} \times \frac{I_i}{I_s} \times 10^{-3} \qquad (2)$$

where X_i represents the content of each compound (μg/kg), V_s denotes the volume of internal standard (μL), Cs signifies the concentration of internal standard (μg/mL), M refers to the content of the tested sample (kg), I_s and I_i are the peak area of internal standard, and compound, respectively.

The relative odor activity value (rOAV) of volatile compounds were a method used to identify the key flavor compounds of food by considering their sensory thresholds. It helps determine the contribution of each aroma compound to the overall aroma profile of the sample. Generally, rOAV >1 indicates that the compound directly contributes to the sample's flavor. The calculation equation is as follows (3):

$$rOAV_i = \frac{C_i}{T_i} \qquad (3)$$

where $rOAV_i$ represents the odor activity value of each compound, C_i denotes the mass concentration of compound (μg/kg), and T_i refers to the odor or aroma threshold of the volatile compound (μg/kg).

2.3. Illumina high-throughput sequencing

Total microbial genomic DNA was extracted using the CTAB method. All PCR reactions were carried out with 15 μL of Phusion® High-Fidelity PCR Master Mix (New England Biolabs). The V3-V4 hypervariable region of bacteria 16S rRNA was amplified using primer pairs 341F (5'-CCTAYGGGRBGCASCAG-3') and 806R (5'-GGACTACHVGGGTWTC-TAAT-3') for 16S rRNA. The fungi ITS1 rDNA were amplified using primer pairs ITS5-1737F (5'-CTTGGTCATTTAGAGGAAGTAA-3') and ITS2-2043R (5'-GCTGCGTTCTTCATCGATGC-3') for ITS. The quantity, purity, and quality of the amplification product were performed using 2 % agarose gels, Qiagen Gel Extraction Kit (Qiagen, Germany), and Qubit@ 2.0 Fluorometer (Thermo Scientific), respectively. The sequencing was conducted on an Illumina Novaseq Platform with PE250 mode (NovaSeq6000 S2 Reagent Kit v1.5, Illumina, USA). Subsequently, the raw reads were filtered using Fastp (v0.22.0, https://github.com/OpenGene/fastp) to obtain clean reads, the clean reads were then spliced using FlASH (v1.2.11, http://ccb.jhu.edu/software/FLASH/), and chimera sequences were removed by UCHIME Algorithm (http://www.drive5.com/usearch/ manual/uchime_algo.html), then the effective tags finally obtained. The amplicon sequence variant (ASV) was generated by Deblur (v1.1.1) based on QIIME2 (v2023.2). In each cluster, the tags with the highest relative abundance were assigned as the representative OTUs. For each representative sequence, the SILVA database (http://www.arb-silva.de/) and UNITE database (https://unite.ut.ee/) were used based on the Mothur algorithm (v1.48) to annotate taxonomic information. The abundance statistics of each taxonomy were visualized using Krona (v2.6). The ASV with a relative abundance of >1 % in each sample were classified as abundant taxa. Alpha diversity indexes were calculated from the representative ASV table via QIIME (v1.9.1) based on the t-test and Wilcoxon rank-sum test. The raw data of Illumina high-throughput sequencing were submitted to the NCBI Short Archive (SRA accession: PRJNA1194492).

2.4. Statistical analysis

Statistical analysis was evaluated by SPSS Statistics software (v26.0; IBM, Armonk, NY, USA) through one-way analysis of variance (ANOVA) along with Duncan's multiple-range test. Principal component analysis (PCA), Orthogonal partial least squares discrimination analysis (OPLS-DA), volcano diagram, Venn diagram, hierarchical cluster analysis (HCA), and flavor wheel were performed using the R package (v3.5.1) at the Metware Cloud (https://cloud.metware.cn). A histogram of metabolic abundance and a boxplot of alpha diversity was conducted using GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA, USA). The co-occurrence network analysis ($|r| \geq 0.9$, $p < 0.05$) was visualized using Gephi (v10.1; Association Gephi, Paris, France). The Mantel test was performed using ggcor 0.9.8.1 based on the R project (v4.2.0). Graphic layouts were generated using Adobe Illustrator 2019 (Adobe Systems Incorporated, San Jose, CA, USA).

3. Results and discussion

3.1. Quality control analysis of samples

Fig. S1 shows the total ion current (TIC) mass spectra obtained using the HS-SPME-GC/MS for quality control (QC) samples in SIM ion detection mode. The x-axis represents the retention time of the metabolite assay, while the y-axis represents the intensity of the ion current of the ion assay. The quality control samples are prepared by mixing all samples and are used to evaluate reproducibility by adopting the same treatment method. During the instrument analysis process, one QC sample is inserted for every ten test samples to monitor the reliability of the analysis. The reproducibility of metabolite extraction and detection is evaluated by overlapping and analyzing the total ion current chromatograms of different QC samples in mass spectrometry detection.

Significant overlap of the total ion current chromatograms of different quality control samples means that metabolite extraction and detection have good reproducibility. The experimental results show that the total ion current spectra of metabolites are highly overlapped, and the retention time and peak intensity between QC samples are consistent. As shown in Fig. S1, the highly overlapped spectra indicate that this method has good signal stability and data reliability.

3.2. The effects of different mucilage retention on volatile compounds based on HS-SPME-GC–MS analysis

3.2.1. Characterization of volatile compounds

Fig. 1A shows the samples with different processing treatments, illustrating a gradual deepening color of the coffee beans as the mucilage retention increases. The volatile compounds of different mucilage retention treatments were analyzed using HS-SPME-GC–MS. A total of 257 volatile compounds were identified in five treatments, including 46 easters (17.90 %), 44 hydrocarbons (17.12 %), 42 heterocyclic compounds (16.34 %), 36 terpenoids (14.01 %), 24 aldehydes (9.34 %), 20 alcohols (7.78 %), 14 ketones (5.45 %), 12 aromatics (4.67 %), and more (Fig. 1B). The classes of identified volatile compounds were consistent with previous studies, which comprises hydrocarbons, alcohols, aldehydes, ketones, acids, esters (Pereira et al., 2019). However, the most

abundant classes were different from those reported in the literature, which may be due to differences in processing methods and fermentation styles. The heatmap of the metabolic abundance of all identified volatile compounds showed that metabolite abundance was significantly higher in the RH group, suggesting that 75 % ~ 80 % mucilage retention led to an increase in aroma compounds (Fig. 2C). In contrast to the other metabolites, the abundance of acids was significantly lower in RH group.

Applying PCA and OPLS-DA multivariate analysis to the HS-SPME-GC–MS matrix data of coffee samples with different mucilage retention treatments. PCA is a common method for reducing dataset dimensionality while retaining the features contributing most to its variance. In the PCA scores plot (Fig. 1D), the cumulative contribution of the first two principal components was 60.71 %, with PC1 and PC2 accounting for 45.34 % and 15.37 %, respectively. The PCA scores plot illustrates a distinct separation between the RH group and the other treatment groups, which exhibit a unique metabolic profile, whereas the differences between the other treatments were not significant. Then, the differential variables were filtered by removing unimportant differences through the OPLS-DA model to obtain a VIP value for each metabolite. The OPLS-DA score diagram (Fig. 2A-D) demonstrated a clear separation between the different treatments (CH vs. WH; WH vs. YH; YH vs. RH; RH vs. BH). The R^2Y and Q^2 values for each model exceeded 0.99 and 0.70,
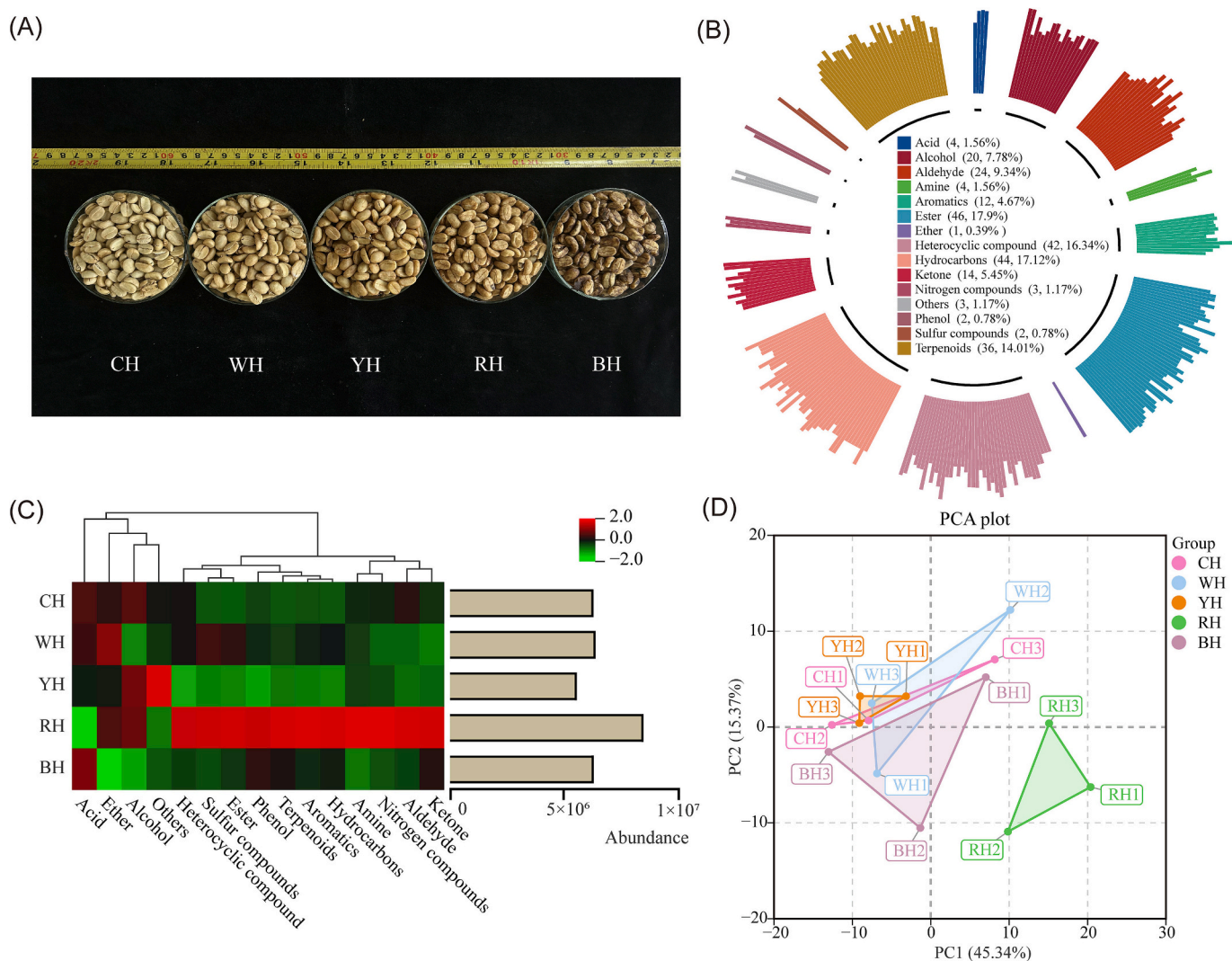

Fig. 1. (A) Coffee samples after different mucilage retention treatments; (B) Proportion circle diagram of metabolite classification; (C) Abundance cluster heatmap of changes in peak area of volatile compound classes at different mucilage retention treatments; (D) PCA analysis of volatile compounds.
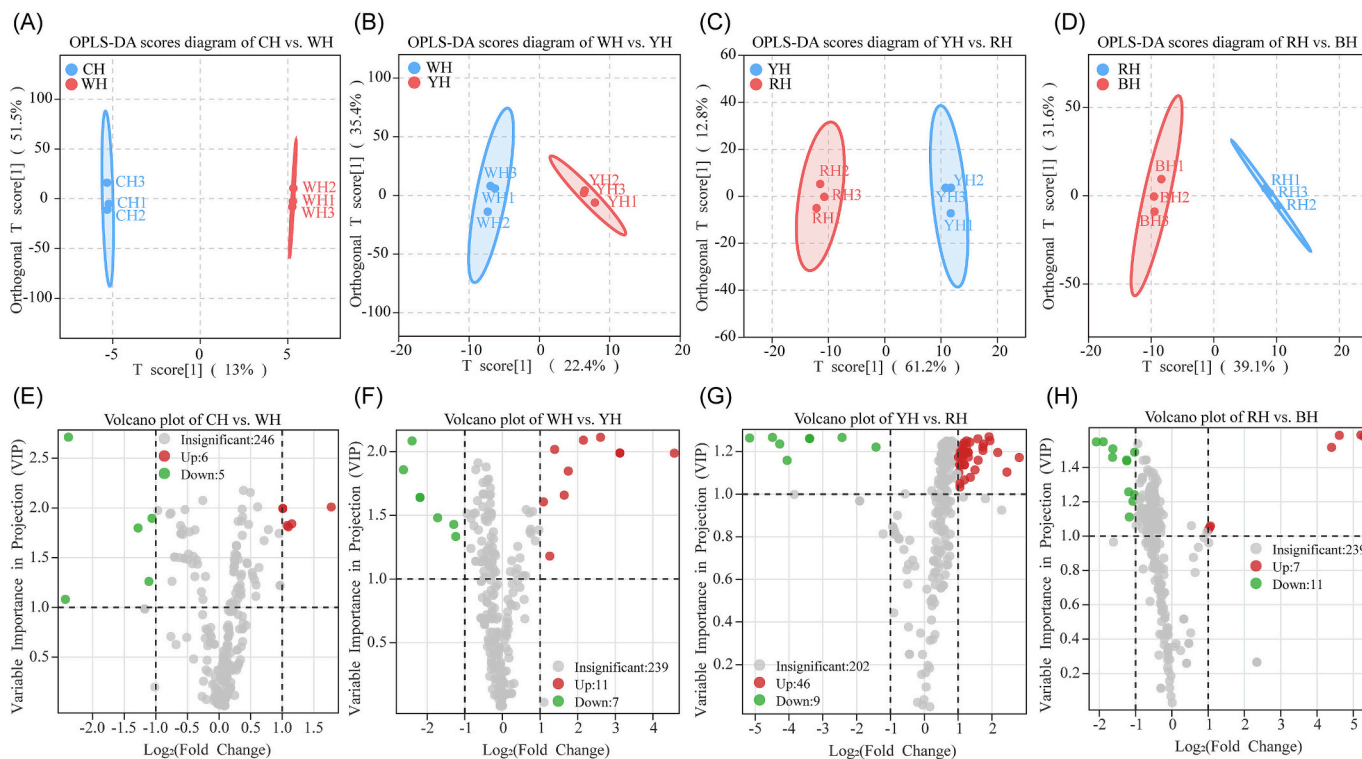
Fig. 2. OPLS-DA analysis of different mucilage retention treatments of coffee beans: (A) CH vs. WH; (B) WH vs. YH; (C) YH vs. RH; (D) RH vs. BH. Volcano plots of differential volatile compounds: (E) CH vs. WH; (F) WH vs. YH; (G) YH vs. RH; (H) RH vs. BH.

respectively, based on the OPLS-DA permutation plots (Fig. S2). In general, A model with $Q^2 > 0.5$ can be considered reliable, while $Q^2 > 0.9$ can be considered excellent (Thévenot et al., 2015). All Q^2 values exceeded 0.7 in this study, indicating that all analyses of the model were statistically acceptable and reliable.

3.2.2. Differential analysis of volatile compounds

Based on the variable importance in projection (VIP) values obtained from the OPLS-DA model and the value of the $\log_2|\text{fold change}| > 1$, we identified 11, 18, 55, and 18 differential metabolites in the comparison groups of CH vs. WH, WH vs. YH, YH vs. RH, and RH vs. BH, respectively. Volcano plots were utilized to illustrate the differences in volatile compounds between the two mucilage retention treatment groups, with larger absolute values on the horizontal axis indicating more significant differences (Fig. 2E-H).

Compared to the CH group, 6 volatile compounds were up-regulated and 5 compounds were down-regulated in the WH group, mainly esters, ketones, and aromatics. Among them, KMW0074 (butanoic acid, ethyl ester) was the most significantly up-regulated volatile compound based on $\log_2|\text{fold change}|$ value. The relative content of KMW0074 in the WH group was significantly higher than that in the CH group and remained stable in different mucilage retention treatments. Typically, medium and long-chain fatty acid ethyl esters such as KMW0074 are one of the primary contributors to the fermentation flavors observed in Baijiu, which impart "fresh" and "fermented" aromas to coffee beans during mucilage fermentation (Xu et al., 2023). The relative content of the KMW0709 (hexadecanoic acid, ethyl ester), a most significantly down-regulated volatile compound, was lower in the WH and YH groups, potentially reducing the mellowness of the coffee flavor.

There were 11 up-regulated and 7 down-regulated volatile compounds in the YH group compared to the WH group, mainly esters, heterocyclic compounds, and aldehydes. The relative content of heterocyclic compounds such as KMW0264*022 (pyrazine, 3-ethyl-2,5-dimethyl-), KMW0269*022 (2,3-dimethyl-5-ethylpyrazine), KMW0268*022 (pyrazine, 2-ethyl-3,5-dimethyl-), which mainly

provides "burnt", "nutty", and "roasted" aromas, increased significantly in the YH group (Liu et al., 2019). Heterocyclic compounds, which provide the distinctive flavor of roasted coffee beans, are not normally found in green coffee beans, except 2-methoxy-3-(2-methylpropyl)-pyrazine (Lee & Shibamoto, 2002). The present study challenges this hypothesis and suggests that the occurrence of heterocyclic compounds in green coffee beans may depend on processing methods. In addition, the relative contents of XMW1309 (pyrazine, tetramethyl-) and D272*145 (butanoic acid, 3-methyl-2-methylbutyl ester) were significantly lower in the YH group. Of these, XMW1309 contributes to the "musty" flavor; D272*145 provides the "herbal" and "earthy" flavor (Lee et al., 2016; Wang et al., 2020). The reduction in the content of these volatile compounds facilitates the formation of high-quality coffee flavors.

The largest number of differential metabolites, comprising 46 up-regulated metabolites and 9 down-regulated metabolites, were identified in the comparison of the YH and RH groups. The most abundant differential metabolites observed at this stage were terpenoids, exhibiting significantly higher relative content in the RH group than the other treatments. Terpenoids are believed to be responsible for the distinct aroma of fruits such as strawberries, citrus, and mango (El Hadi et al., 2013). Moreover, terpenoids such as linalool and α-pinoresinol have been demonstrated to possess antibacterial and anti-inflammatory properties (Park et al., 2012). Apart from homoterpenes, the plant produces other irregular volatile terpenoids such as β-ionone and β-damascenone, which are usually produced by the degradation of carotenoids and can produce significant "floral" and "fruity" aromas (Dudareva et al., 2006). In addition, esters such as KMW0709 (hexadecanoic acid, ethyl ester), KMW0204 (acetic acid, hexyl ester), and D202 (3-hexen-1-ol, propanoate, (Z)-) were significantly increased in the RH group, giving the coffee beans a "mellow" flavor and a variety of "fruity" aromas.

In the comparison between the RH and BH groups, 7 up-regulated and 11 down-regulated volatile compounds were identified, mainly esters, heterocyclic compounds, and ketones. In the BH group, the relative content of the majority of the esters that control flavors were

diminished, which resulted in a reduction of the "fresh", "rose", and a variety of "fruity" aromas present in the coffee beans. However, the relative content of KMW0716 (isopropyl-palmitate) was significantly increased in this treatment, which might improve the aroma of the "oils" in the coffee beans.

A total of 85 differentially volatile compounds were identified in the various comparison groups, with some exhibiting differential expression across the comparison of multiple groups (Fig. 3A, Fig. S3). Among these compounds, only KMW0441 (benzeneacetic acid, ethyl ester) showed significant differential expression across four comparison groups, with its relative content being significantly higher in the WH and RH groups compared to the other treatments. KMW0441 may be related to the shikimate pathway and is usually formed during fermentation through the esterification of phenylacetic acid following its interaction with microorganisms (Tat et al., 2007). KMW0441 reportedly produces "sweet" and "minty" flavors and substantially enhances overall aroma quality, which may benefit the coffee's overall aroma contribution (Fan et al., 2020; Hou et al., 2022). A comparison of the groups revealed that YH vs. RH exhibited the highest number of differential volatile compounds (Fig. 3A). Furthermore, with regard to the peak area of all differential metabolites, the RH group was the most abundant compared to the other treatments (Fig. 3B). The significant differences in the aroma profiles of these volatile compounds may cause coffee beans with different mucilage retention treatments to produce different flavors. This result indicates that quality control is essential during the honey processing of coffee beans.

3.2.3. Flavor wheel of differential volatile compounds

Not all of the volatile compounds contribute significantly to flavor development. In general, the relative content of volatile compounds exceeds their thresholds to be considered a significant contributor to aroma. Nevertheless, some volatile compounds with specific aromatic profiles can be used in combinations to produce particular flavors (Liang et al., 2024). Each differential volatile metabolite was identified by consulting the website (https://www.thegoodscentscompany.com/index.html) and relevant papers for its aroma attributes. For each comparison group, flavor wheels were created based on the differential metabolites and their annotated flavor profiles, and the 10 flavor profiles with the highest number of annotations were selected. As shown in Fig. S4, "sweet, fresh, fruity" were the primary flavor profiles in the CH vs. WH group, while "nutty, cocoa, burnt, roasted" were the major flavor profiles in the WH vs. YH group. The differential volatile compounds with flavor profile annotations were mostly up-regulated in the YH vs. RH group. A total of 61 flavor profiles were identified in these compounds, with annotations such as "fruity, sweet, green, floral, woody"

being attributed to up to 5 differential compounds. In the RH vs. BH group, the predominant flavor profiles were identified as "fruity, banana, green, sweet, apple." However, the differential volatile compounds affecting these high-quality flavor profiles were generally down-regulated, which may negatively affect coffee flavor.

Generally, "sweet", "fruity", and "green" flavors are common in fermented coffees, including those that undergo mucilage fermentation and exogenously added yeast (Pereira et al., 2024). These aroma and flavor precursors formed during fermentation ultimately influence the flavor formation of coffee roasted beans by migrating into the coffee green beans (Ferreira et al., 2023). Surprisingly, coffee beans processed with 75 % ~ 80 % mucilage retention (RH) seem to promote "floral" sensations, an essential flavor in coffee that is highly appreciated in the market and can markedly elevate the selling price. This flavor is associated with the variety and processing method and is typically found in "Geisha", which is known for the "floral" sensations (Marie et al., 2024). Additionally, the exogenous addition of yeast has been reported to produce "floral" sensations (Pereira et al., 2024; Tang et al., 2021). Therefore, further studies are required to validate the effect of fermentation control on floral sensations to enhance the attractiveness and economic value of coffee products.

3.2.4. Screening of key aroma compounds by rOAVs

To identify the key aroma compounds affecting the overall flavor of the coffee following different mucilage retention treatments, rOAVs of all differential volatile compounds were calculated. rOAVs are frequently applied to evaluate the contributions of aroma compounds. The compounds with an rOAV > 1 are considered significant contributors to the aroma (Huang et al., 2022). In this study, a total of 32 metabolite thresholds were found for 85 differential volatile compounds. Of these, 13 volatile compounds with an rOAV >1 were considered key aroma compounds affecting coffee flavor with different mucilage retention treatments (Table 1). These key aroma compounds include 4 esters, 3 aromatics, 3 terpenoids, 2 heterocyclic compounds, and 1 ketone.

The rOAVs of KMW0230 (3-octen-2-one), KMW0074 (butanoic acid, ethyl ester), and KMW0094 (butanoic acid, 3-methyl, ethyl ester), in the green coffee bean samples, were notably high (average rOAV of different treatments exceeding 100), which may be attributed to the low thresholds of 0.03, 0.053, and 0.1, respectively. These esters and ketones may serve as the basis for the flavor profile of coffee. Among them, as mentioned earlier, KMW0074 brings out "fresh" and "alcoholic" flavors, with an rOAV significantly higher in all coffee beans that have not completely mucilage removal, which can be said to be caused by mucilage fermentation. KMW0230 produced coffee with "earthy, spicy,
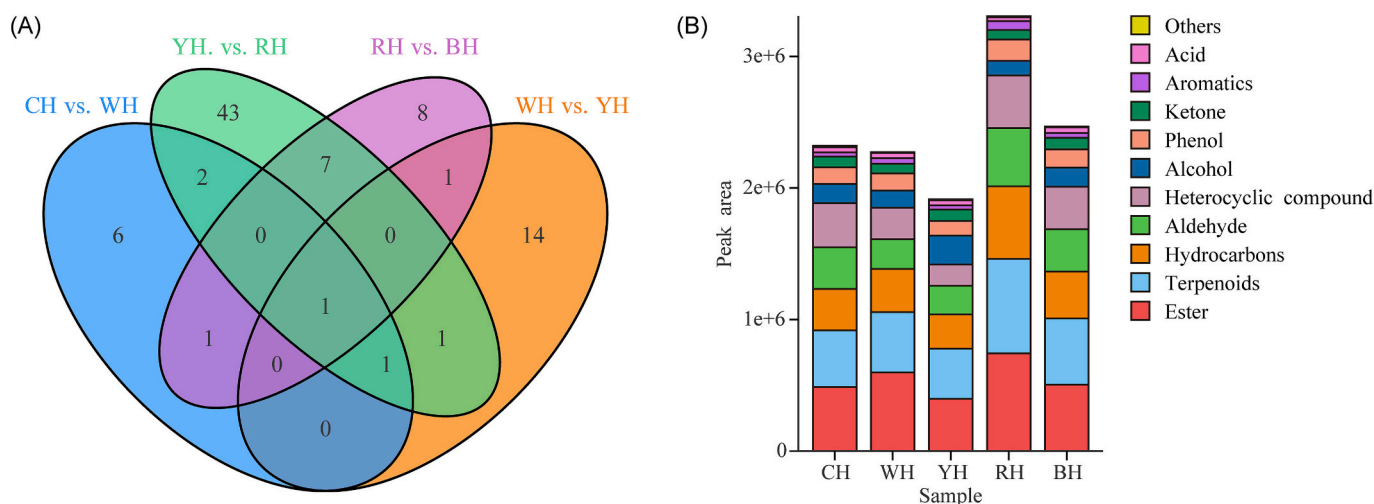

Fig. 3. (A) Venn diagram of differential volatile compounds for each comparison group; (B) Histogram of relative content of differential volatile compounds.

Table 1
The key aroma compounds with high odor activity values (rOAVs).

No	Compound	Class	Odor profiles	Relative content (μg/kg)					Threshold (μg/kg)	rOAV				
				CH	WH	YH	RH	BH		CH	WH	YH	RH	BH
1	naphthalene, 1-methyl-	Aromatics	naphthyl, chemical, medicinal, camphor	3.22	8.44	5.77	7.50	4.78	8	0.40	1.06	0.72	0.94	0.60
2	naphthalene, 2-methyl-	Aromatics	sweet, floral, woody	3.22	8.44	5.77	7.50	4.78	4	0.80	2.11	1.44	1.88	1.19
3	styrene	Aromatics	penetrating, balsamic, gasoline	7.77	9.41	7.00	16.89	7.43	3.6	2.16	2.62	1.94	4.69	2.06
4	butanoic acid, ethyl ester	Ester	fresh, alcoholic	4.77	21.38	20.65	9.81	18.64	0.053	90.05	403.35	389.61	185.15	351.64
5	1-ethylpropyl acetate	Ester	–	18.09	21.38	6.38	25.84	18.64	9	2.01	2.38	0.71	2.87	2.07
6	benzoic acid, methyl ester	Ester	phenol, wintergreen, almond, floral, canga	0.64	1.12	0.66	1.37	1.05	0.52	1.22	2.15	1.27	2.63	2.02
7	butanoic acid, 3-methyl-, ethyl ester	Ester	strawberry, candy, fruity	18.09	21.38	20.65	33.40	18.64	0.1	180.94	213.77	206.50	333.96	186.37
8	pyrazine, 2-ethyl-3,5-dimethyl-	Heterocyclic compound	burnt, almond, roasted, nutty, coffee	0.77	0.46	3.73	0.82	1.07	0.04	19.21	11.42	93.14	20.39	26.81
9	2(5H)-furanone, 5-ethyl-	Heterocyclic compound	spice	114.74	83.24	53.24	139.81	117.35	9.7	11.83	8.58	5.49	14.41	12.10
10	3-octen-2-one	Ketone	earthy, spicy, herbal, sweet, mushroom, hay, blueberry	15.86	13.26	11.31	15.95	8.07	0.03	528.73	442.06	377.11	531.52	268.95
11	trans-β-ionone	Terpenoids	dry, powdery, floral, woody, orris	9.08	11.20	9.90	16.84	10.25	0.2	45.38	56.02	49.49	84.19	51.24
12	β-ionone	Terpenoids	floral, woody, sweet, fruity, berry, tropical, beeswax	9.08	11.20	9.90	16.84	10.25	0.12	75.64	93.37	82.48	140.32	85.41
13	linalool	Terpenoids	floral, green	3.64	5.40	3.95	6.57	4.81	6	0.60	0.90	0.66	1.09	0.80

herbal, sweet, mushroom, hay, and blueberry" aromas, with significantly lower content in the BH group. The relative content of KMW0094 was significantly higher in the RH group, imparting "strawberry", "candy", and "fruity" aromas to coffee. Furthermore, KMW0267 (benzoic acid, methyl ester), gives coffee "phenol, wintergreen, almond, floral, and canga" aroma, and has only been detected in fermented green coffee beans (Wan, Li, et al., 2024). The KMW0268*022 (pyrazine, 2-ethyl-3,5-dimethyl-) and KMW0539 (2(5H)-furanone, 5-ethyl-) are heterocyclic compounds with rOAV >1, exhibiting predominantly "burnt, almond, roasted, nutty, coffee, and spice" aroma. As previously described, heterocyclic compounds were typically present in roasted beans, imparting a pronounced roasted flavor. The high relative content of the YH group may contribute to this flavor enhancement.

Terpenoids, such as β-ionone, trans-β-ionone, and linalool, impart a distinct "floral" flavor to coffee. These aroma compounds in question have been identified in tea and roasted coffee beans as the primary substances responsible for imparting a "floral" flavor (Liang et al., 2024; Zhai, Dong, Fu, et al., 2024; Zhu et al., 2018). The rOAV of β-ionone and trans-β-ionone in the RH group, respectively, was 140.32 and 84.19, which was the highest among all the treatments, while linalool had rOAV >1 only in the RH group with a threshold of 6. These results suggested that β-ionone, trans-β-ionone, and linalool were the predominant contributors to the floral sensation of the coffee samples, and the RH group exhibited the most pronounced floral sensation. During roasting, these compounds underwent transformations, leading to the development of more intricate ingredients that contributed to the overall flavors and aromas (Wan, Li, et al., 2024).

Aromatics were the most important natural aromatic compounds.

Studies have shown that they were particularly abundant in green coffee beans of the "Catimor" variety (Wan, Wang, et al., 2024). The content of the aromatics including w08*082 (naphthalene, 1-methyl-), XMW0662*082 (naphthalene, 2-methyl-), and styrene increased significantly through the mucilage fermentation process. For example, naphthalene, 1-methyl- with an rOAV >1 only in the WH group produced a "naphthyl, chemical, medicinal, and camphor" odor that can severely affect the flavor of coffee. In contrast, the presence of naphthalene, 2-methyl- helps coffee to produce "sweet", "floral", and "woody" aroma. It has been found that naphthalene, 2-methyl- contributes greatly to the refreshing aroma of tea (Wang, Qu, et al., 2022). Styrene has a high content in the RH group, imparting "penetrating", "balsamic", and "gasoline" aromas. This may increase the mellowness of the coffee, but in excess, it can lead to unpleasant characteristics. However, these compounds have relatively high odor thresholds, which may result in an insignificant overall impact on the aroma (Wang, Shi, et al., 2022). In addition, terpenoids and aromatics are among the antioxidants naturally found in coffee beans and may help to minimize the production of heat hazards such as acrylamide (Kahkeshani et al., 2015).

3.3. Differences in microbial communities during different mucilage retention fermentation

3.3.1. Richness and diversity of bacteria and fungi communities

A total of 306,8034 and 110,3804 effective tags were generated from 16S and ITS genes, respectively, from thirty green coffee bean samples (six replicates per group). The mean lengths of the effective tags for 16S and ITS were 406 bp (a range from 406 bp to 408 bp) and 267 bp (a

range from 232 bp to 312 bp). Clustering of non-repetitive sequences with 97 % similarity identified 349 bacterial ASVs and 1092 fungal ASVs, respectively. Subsequently, the richness and diversity of microbial communities were assessed using the Chao1, ACE, Shannon, and Simpson indices of alpha diversity. Among them, Chao1 and ACE are commonly used to assess the species richness of samples, while Shannon and Simpson are used to assess the evenness and diversity of samples. The results demonstrated that the Chao1 and ACE indices of the bacteria communities reached the highest level in the WH group, followed by a continuous decrease. A similar trend was observed in the Shannon and Simpson indices, with the exception that the lowest level was exhibited in the YH group(Fig. 4A). The Chao1, ACE, Shannon, and Simpson indices of fungi communities exhibited a decreasing and then increasing trend, reaching the lowest level in the YH group(Fig. 4B). Overall, the trend indicated an increase in bacterial diversity and species richness with increasing mucilage retention, followed by a subsequent decrease. In contrast, the opposite trend was observed for fungi.

3.3.2. Taxonomic characteristics of bacteria communities

Through alignment and annotation, 15 phyla and 63 genera of bacterial communities were identified in five treatments. With >1 % abundance as the screening condition, the dominant phyla included Cyanobacteria, Proteobacteria, and Firmicutes (Fig. 5A), consistent with previously reported results (Shen et al., 2024). At the genus level, the dominant genera were *Kosakonia* and *unclassified_f_Enterobacteriaceae*, followed by *Weissella*, *Pediococcus*, *Lactococcus*, *Pantoea*, and *Lactiplantibacillus*, with the relative abundance >1 % (Fig. 5B). As the mucilage retention on the coffee beans increased, the abundance of most of the bacterial genera (such as *Kosakonia*, *unclassified_f_Enterobacteriaceae*, *Weissella*, *Pediococcus*, *Lactococcus*, and *Lactiplantibacillus*) initially increased and then decreased, exhibiting the highest in the WH group. This phenomenon may be due to the group with higher mucilage retention requiring longer for natural fermentation to the specified moisture content. In the early stages of fermentation, lactic acid bacteria, such as *Lactiplantibacillus*, *Lactococcus*, and *Wessella* were prevalent

and induced a decreased pH by producing organic acids through continuous lactic acid fermentation. As fermentation progresses, mixed-acid bacteria (*Enterobacteriaceae*, *Kosakonia*) tend to homogenize with LAB, which may be related to environmental acidification. Studies have shown that microbial dynamics are gradually becoming homogenized during processing, despite differences in processing methods, consistent with our study (Pereira et al., 2024). This is because high acidity reduces bacterial diversity and ensures the presence of acid-tolerant microorganisms (Cruz-O'Byrne et al., 2021). Furthermore, the RH group exhibited a uniquely high abundance of *Pantoea* (18.32 ± 2.41 %). *Pantoea* is renowned for its abundant metabolism ability, which facilitates the breakdown of carbohydrates and compound synthesis such as alcohols and esters (Vale et al., 2024). Its increasing trend may be attributed to the late effect of mucilage fermentation, which could potentially influence the ultimate flavor profile of the coffee bean.

3.3.3. Taxonomic characteristics of fungi communities

A total of 4 phyla and 110 genera were identified in the five treatments with different mucilage retention levels. At the phylum level, only two fungal phyla exhibited relative abundance >1 %, namely Ascomycota and Basidiomycota (Fig. 5C). Of these, Ascomycota is the dominant fungal phylum, with an abundance exceeding 90 % in all honey processing method. The abundance of Ascomycota in correlation increased with mounting mucus retention, peaking in the RH group (99.53 ± 0.05 %) and decreasing to 90.07 ± 2.00 % in the BH group. The opposite trend was observed for Basidiomycota, reaching a minimum of 0.35 ± 0.05 % in the RH group and subsequently increasing to 7.88 ± 2.23 % in the BH group. The fungal phyla Ascomycota and Basidiomycota have been identified as key players in the fermentation of coffee, tea, and white wine, contributing significantly to the quality and physicochemical properties of the fermented products (Fu et al., 2024; Zheng et al., 2022; Zhou et al., 2024).

At the genus level, the dominant genera were *unclassified_f_Nectriaceae*, *unclassified_O_Saccharomycetales*, *Hanseniaspora*, and *Cladosporium*, with a relative abundance >1 % (Fig. 5D). Among them,
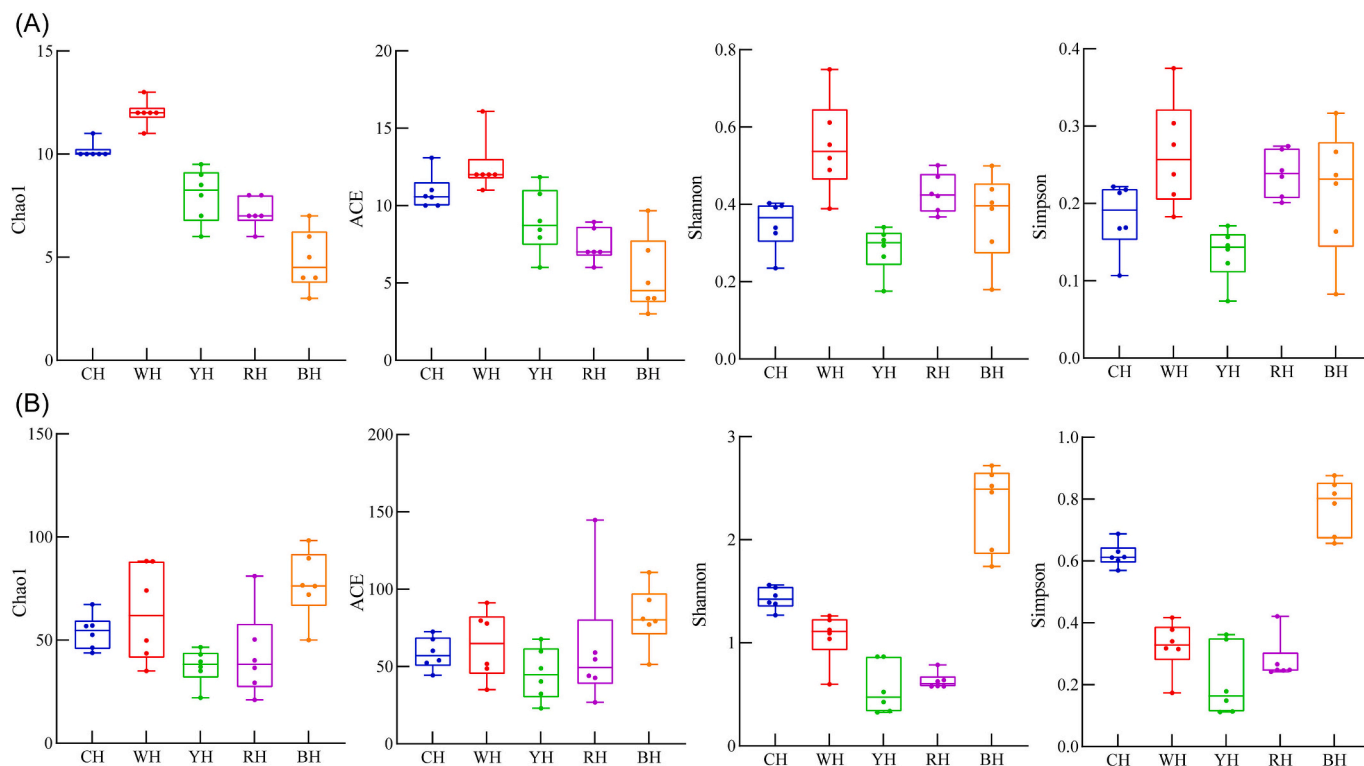

Fig. 4. Richness and diversity of microbial communities (A. bacteria; B. fungi) in different mucilage retention treatments of coffee beans.
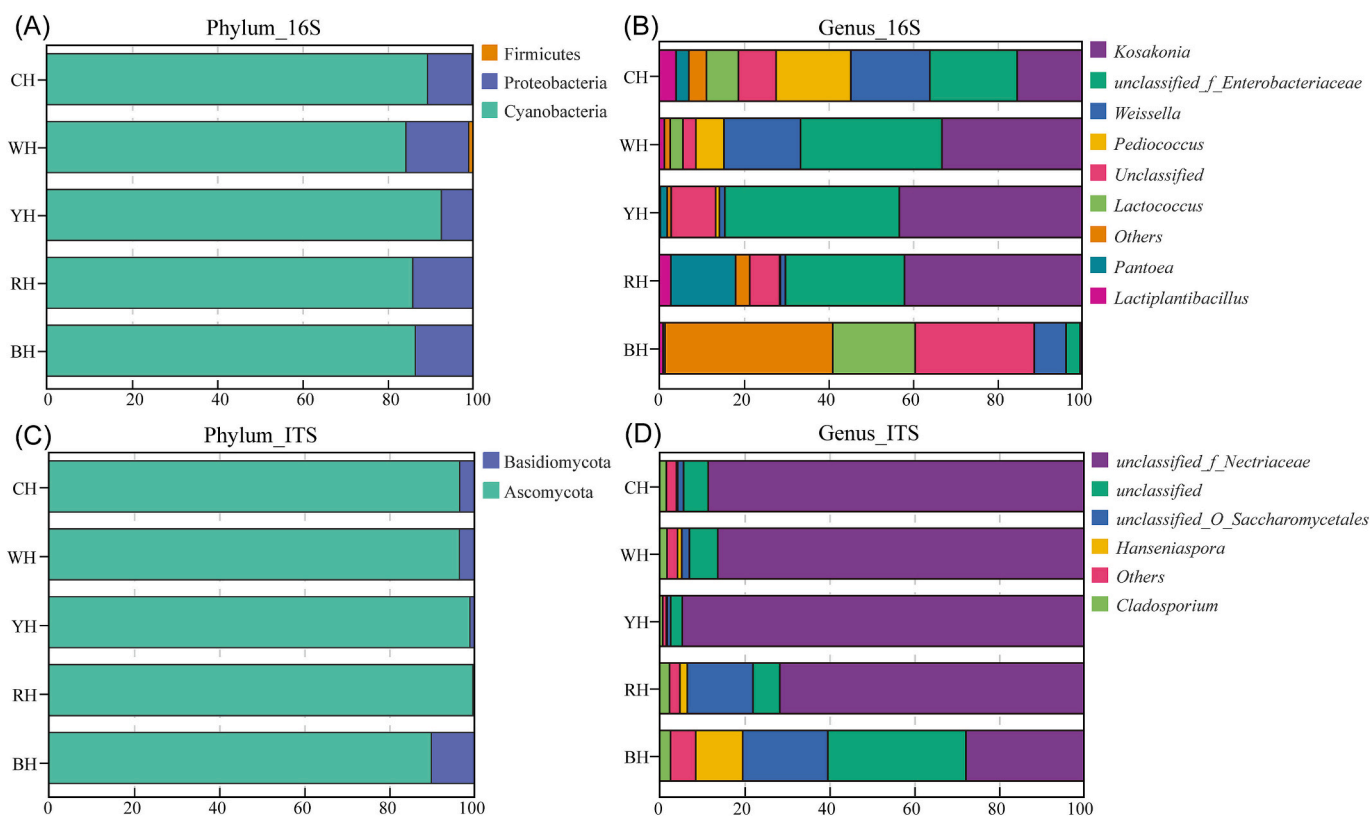
Fig. 5. (A) Bacterial communities structure bar plot in phylum level; (B) Bacterial communities structure bar plot in genus level; (C) Fungal communities structure bar plot in phylum level; (D) Fungal communities structure bar plot in genus level.

unclassified fungi belonging to the family Nectriaceae and the genus *Cladosporium*, both filamentous fungi, had an average abundance of 70.67 ± 5.07 %, and 1.95 ± 0.18 %, respectively, for different honey processing methods. Filamentous fungi exhibited high abundance across all treatments with different mucus retention, while the relative abundance of *unclassified_f_Nectriaceae* demonstrated a continued decrease in the RH and BH groups. The reason for this phenomenon may be attributed to the long natural fermentation time and high acidity of high mucus retention. This is feasible because filamentous fungi may have the potential to produce toxins (Bravim et al., 2023). Nevertheless, research has demonstrated that these filamentous fungi remain an integral component of the coffee microbiota (Pregolini et al., 2021; Vale et al., 2024). Observations of the family Nectriaceae, genus *Cladosporium*, and even genus *Colletotrichum* during coffee fermentation in numerous regions have been observed, such as Brazil, Australia, and Honduras. While these communities may facilitate mucilage breakdown, the precise function remains uncertain. The unclassified fungi of the Order Saccharomycetales and genus *Hanseniaspora*, both of which are yeasts, were significantly more abundant in the RH and BH groups with high mucilage retention. The relative abundance of *unclassified_o_Saccharomycetales* was 1.00 ± 0.62 %, 17.70 ± 2.33 %, and 27.65 ± 5.32 % in YH, RH, and BH, respectively; and that of *Hanseniaspora* was 0.32 ± 0.19 %, 2.41 ± 2.93 %, and 22.86 ± 11.80 % in YH, RH, and BH, respectively. The trend in abundance of these non-Saccharomyces yeasts is opposite to that of mixed-acid bacteria and lactic acid bacteria, as we observed in the alpha diversity (Fig. 4). It has been demonstrated that non-Saccharomyces yeasts exhibit a significant correlation with ethanol production (Pereira et al., 2024). Due to their complex metabolic process producing aromatic compounds, non-Saccharomyces yeasts play a vital role in coffee fermentation (Dzialo et al., 2017). These aromatic compounds may contribute to the regulation of microbial community signaling or the release of aroma compounds (e.g. terpenes, alcohols, esters, aldehydes, etc.) through enzymatic hydrolysis of glycosidic

precursors as well as metabolic pathways.

3.4. Microbial association with volatile compound in the different mucilage retention treatments

3.4.1. Co-occurrence network analysis of microbial genera and differential volatile compounds

Microorganisms are considered to be one of the sources of metabolites. Many studies have established that the presence of many yeasts and bacterial communities during fermentation can modulate the flavor profiles of coffee beverages (Pereira et al., 2015; Silva et al., 2013). In the present study, the correlation between dominant microbial (bacteria and fungi) communities and differential volatile compounds was initially established through co-occurrence network analysis. Using criteria of $|r| \geq 0.9$ and $p < 0.05$, 7 dominant bacterial genera were significantly associated with 12 metabolites, while 4 dominant fungal genera were significantly linked to 41 metabolites, as illustrated in Fig. 6A. For mixed acid bacteria, three bacterial genera including *Pediococcus*, *Kosakonia*, and *unclassified_f_Enterobacteriaceae* were significantly negatively correlated with only three metabolites, KMW0078 (8-nonen-2-one), XMW0855 (ethanone-1-(2-methyl-1,3-dithiolan-2-yl)-), and KMW0373 (benzoic acid, ethyl ester); whereas *Pantoea* exhibited a significantly positively correlation with four metabolites, WMW0216 (fenchyl acetate), XMW1286 (1H-imidazole-4-methanol), XMW0499 (undecane-3,4-dimethyl-), and XMW1440 (2-methylprop-2-enoyl 2-methylprop-2-enoate). *Pantoea* has been reported in previous studies as a common bacteria genus involved in the natural fermentation of coffee, which participates in the fermentation process by degrading mucilage via the mixed acid pathway (Góngora et al., 2024). Our study suggests that it may also contribute to the synthesis of flavor compounds, especially in specific retention treatments. In contrast, lactic acid bacteria (*Lactiplantibacillus*, *Lactococcus*, and *Weissella*) showed a significantly positive correlation with 3 esters (XMW0073*145 (3-
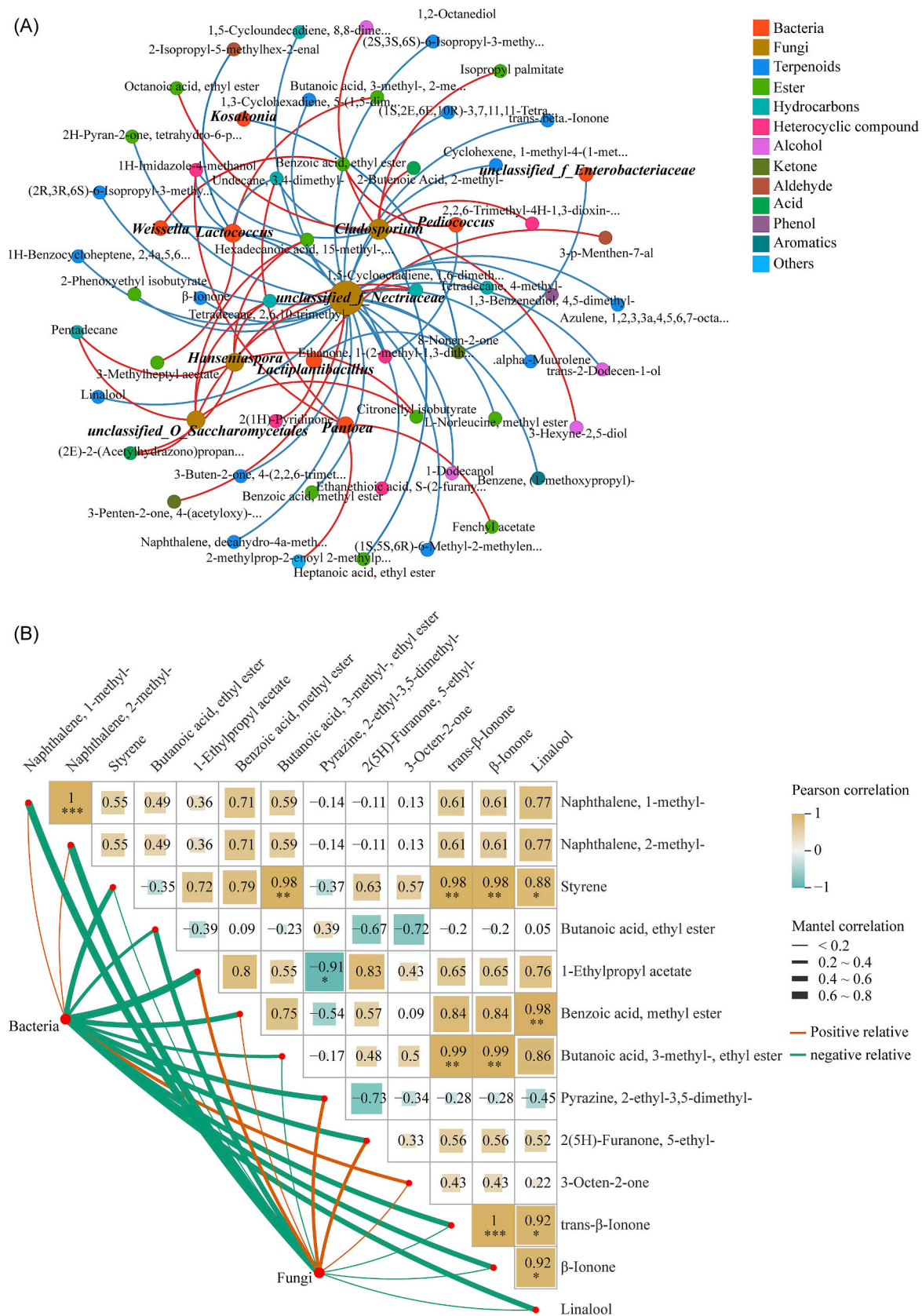
Fig. 6. (A) Co-occurrence network analysis between microbial communities and differential volatile compounds; (B) Mantel test for the associations of microbial communities with 13 key aroma compounds.

methylheptyl acetate), KMW0373 (benzoic acid,ethyl ester), and D272*145 (butanoic acid-3-methyl-2-methylbutyl ester)) to bring out the flavor of "fruit" and "sweet". Esters are widely known to contribute to floral and fruity sensory notes in alcoholic beverages (Procopio et al., 2011). The metabolism of lactic acid bacteria has been shown to contribute to the formation of esters such as ethyl acetate, isoamyl acetate, propyl acetate, ethyl hexanoate, and n-butyl acetate, which contribute to the development of tropical flavor in coffee beverages (Pereira et al., 2015). Nevertheless, the specific contribution of lactic acid bacteria to coffee fermentation remains uncertain. Following the attainment of a specified moisture content in green coffee beans, the numbers of lactic acid bacteria are typically low at this juncture (Elhalis et al., 2023). Our research also identified this phenomenon, notably, the highest abundance of *Lactiplantibacillus*, *Lactococcus*, and *Weissella* was observed in the WH group with lower mucilage retention. This finding warrants further control of the fermentation process and investigating the underlying causes.

For fungal genera, *unclassified_o_Saccharomycetales* and *Hanseniaspora* were significantly positively related with 6 metabolites, including 3 hydrocarbons (XMW0265, XMW0437, and KMW0602), 2 esters (XMW0757 and KMW0585), and 1 acid (XMW1234). This suggests a synergistic effect of producing volatile compounds existed between *unclassified_o_Saccharomycetales* and *Hanseniaspora*. The research result also confirmed that these non-Saccharomyces yeasts contribute to the overall aroma profile development by producing higher esters, acids, and hydrocarbons (Chen et al., 2022). Uniquely, *unclassified_f_Nectriaceae* was significantly negatively related with 26 metabolites, including 14 terpenoids (KMW0595*084, XMW0157*063, XMW0216*063, WMW0051*150, KMW0583*150, KMW0291, XMW0504*084, KMW0606, KMW0604, XMW0544, XMW0817, KMW0556, KMW0639, and KMW0296), 4 esters (KMW0568, KMW0267, D158, and XMW0575), 2 alcohols (KMW0629 and KMW0674), 2 hydrocarbons (XMW0200 and XMW0097), 1 aromatic (XMW0726), 1 heterocyclic compound (KMW0368), and 1 phenol (XMW0455). The correspondence between index and compound was shown in Table S1. Of particular interest was the significant association observed between *unclassified_f_Nectriaceae* and various terpenoids. Nectriaceae are usually described as pathogens in the natural ecosystems. However, some species of *Nectriaceae*, such as *Bisifusarium domesticum*, have been reported as exhibiting potential antimicrobial activity and lipid production capacity (Savary et al., 2023). As a dominant fungus identified in this study, we suggest that some species in *unclassified_f_Nectriaceae* may enhance the flavor quality and bioactivity of coffee beans by affecting the relative content of terpenoids, which requires further investigation. In general, the observed pattern of changes in associated metabolites indicated that bacterial genera primarily functioned in the low mucilage retention treatments (like CH and WH), whereas fungal genera were more active in the high mucilage retention treatments (such as RH and BH).

3.4.2. Relationship between microorganisms and key aroma compounds

To examine the relationship between microorganisms and key aroma compounds in honey processing methods with different mucilage retention, Mantel's test analysis was conducted to map the correlation between dominant bacterial and fungal genera and aroma compounds with rOAV >1 (Fig. 6B). The 11 dominant microbial genera can be categorized into mixed acid bacteria (*Pantoea*, *Kosakonia*, *Pediococcus*, *unclassified_f_Enterobacteriaceae*), lactic acid bacteria (*Weissella*, *Lactococcus*, *Lactiplantibacillus*), filamentous fungi (*unclassified_f_Nectriaceae*, *Cladosporium*), and non-Saccharomyces yeasts (*unclassified_O_Saccharomycetales*, *Hanseniaspora*). The results showed that lactic acid bacteria was significantly positively correlated with naphthalene, 1-methyl- and naphthalene, 2-methyl- ($p < 0.05$). Both compounds were significantly increased in CH vs. WH, suggesting that lactic acid bacteria may act more in treatments with low mucilage retention. In addition, filamentous fungi were significantly proportional to pyrazine, 2-ethyl-3,5-

dimethyl- ($p < 0.05$), and non-Saccharomyces yeasts showed significant positive correlations with three different compounds ($p < 0.05$), namely styrene, benzoic acid, methyl ester, and linalool, respectively. Thus, filamentous fungi may contribute to aromas described as "burnt", "almond", "roasted", "nutty", and "coffee"; whereas non-Saccharomyces yeasts add "mellow", "floral", and "nutty" aromas, which are more prevalent in the RH group. In contrast, the mixed acid bacteria have no significant correlations ($p > 0.05$). The results of the study showed that of all significant correlations, microbial taxa were positively correlated with key aroma compounds, indicating that microbes as a whole contribute to the formation of these compounds, although this correlation was relatively low. Fungi had a greater association with aroma compounds compared to bacteria. Both filamentous fungi and non-Saccharomyces yeasts played significant roles in the formation and accumulation of flavor compounds, particularly observed in the RH group. Therefore, maintaining a mucilage retention of 75–80 % is optimal for the flavor quality and bioactivity of green coffee beans. Of course, further compositional and sensory analysis of baked coffee beans with different mucilage retention treatments is necessary.

4. Conclusion

This study utilized a combination of HS-SPME-GC–MS and high-throughput sequencing to conduct a comprehensive examination of the effects of diverse honey processing methods with different mucilage retention on the volatile compounds and the characteristics of the microbial community in green coffee beans. The HS-SPME-GC–MS result indicated that the differential metabolites varied across different mucilage retention treatments, with the highest abundance of volatile compounds observed at 75–80 % mucilage retention. Based on rOAV >1, 13 key volatile compounds were identified in the honey processing methods, including naphthalene-1-methyl-, naphthalene-2-methyl-, styrene, butanoic acid-ethyl ester, 1-ethylpropyl acetate, benzoic acid-methyl ester, butanoic acid-3-methyl-ethyl ester, pyrazine-2-ethyl-3,5-dimethyl-, 2(5H)-furanone-5-ethyl-, 3-octen-2-one, trans-β- ionone, β-ionone, and linalool, which contribute to excellent flavors described as "mellow", "fruity", "floral", "roasted", and "nutty". The high-throughput sequencing result showed that the diversity of fungal communities exceeded that of bacteria in the honey processing method fermented with different mucilage retention levels. Among them, the dominant bacterial genera were *unclassified_f_Enterobacteriaceae*, *Kosakonia*, *Weissella*, *Pediococcus*, *Lactococcus*, *Pantoea*, and *Lactiplantibacillus*; while *unclassified_O_Saccharomycetales*, *Hanseniaspora*, *Cladosporium*, and *unclassified_f_Nectriaceae* were the dominant fungal genera. The correlation analysis showed that bacteria and fungi contribute to the formation of key aroma compounds. In contrast, bacterial genera functioned primarily in low mucilage retention treatments, whereas fungal genera were more active in high mucilage retention treatments. There is a close association between *unclassified_f_Nectriaceae* and terpenoids. Non-Saccharomyces yeasts, *unclassified_O_Saccharomycetales*, and *Hanseniaspora* were the main genera that promote the formation of aroma compounds. This study aimed to gain further insight into the interactions between flavor compounds and microbial communities by controlling mucilage retention in coffee beans for precise fermentation. The findings of this study provide a theoretical foundation for further insight into the effects of mucilage removal on volatile kinetics in coffee. This knowledge can help coffee producers optimize parameters to produce coffee products with enhanced flavor consistency.

CRediT authorship contribution statement

Faguang Hu: Validation, Resources, Investigation, Funding acquisition, Conceptualization. **Haohao Yu:** Writing – review & editing, Writing – original draft, Formal analysis. **Xingfei Fu:** Validation, Investigation, Formal analysis. **Zhongxian Li:** Data curation. **Wenjiang Dong:** Supervision. **Guiping Li:** Resources. **Yanan Li:** Data curation.

Yaqi Li: Data curation. **Bingqing Qu:** Investigation. **Xiaofei Bi:** Visualization, Supervision, Data curation.

Declaration of competing interest

The authors declare no conflict of interest

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2025.102251.

Data availability

The 16S and ITS rRNA gene sequence data of all samples were submitted to the NCBI SAR database under the accession PRJNA1194492. HS-SPME-GC-MS data will be made available upon request.

References

Aswathi, K. N., Shankar, S. R., Seenivasan, K., Prakash, I., & Murthy, P. S. (2022). Metagenomics and metabolomic profiles of *Coffea canephora* processed by honey/pulped natural technique. *Innovative Food Science & Emerging Technologies, 79*, 12. https://doi.org/10.1016/j.ifset.2022.103058

Aswathi, K. N., Shirke, A., Praveen, A., Chaudhari, S. R., & Murthy, P. S. (2023). Pulped natural/honey robusta coffee fermentation metabolites, physico-chemical and sensory profiles. *Food Chemistry, 429*, 12. https://doi.org/10.1016/j.foodchem.2023.136897

Aswathi, K. N., Shirke, A., Praveen, A., & Murthy, P. S. (2024). Functioning of Saccharomyces cerevisiae in honey coffee (*Coffea canephora*) and their effect on metabolites, volatiles and flavor profiles. *Food Research International, 180*, 10. https://doi.org/10.1016/j.foodres.2024.114092

Bi, X. F., Yu, H. H., Hu, F. G., Fu, X. F., Li, Y. A., Li, Y. Q., & Dong, W. J. (2024). A systematic analysis of the correlation between flavor active differential metabolites and multiple bean ripening stages of *Coffea arabica* L. *Molecules, 29*(1), 13. https://doi.org/10.3390/molecules29010180

Bravim, D. G., de Oliveira, T. M., do Rosário, D. K. A., Batista, N. N., Schwan, R. F., Coelho, J. M., & Bernardes, P. C. (2023). Inoculation of yeast and bacterium in wet-processed *Coffea canephora*. *Food Chemistry, 400*, 9. https://doi.org/10.1016/j.foodchem.2022.134107

Cardoso, W. S., Agnoletti, B. Z., de Freitas, R., de Abreu Pinheiro, F., & Pereira, L. L. (2021). Biochemical aspects of coffee fermentation. In L. Louzada Pereira, & T. Rizzo Moreira (Eds.), *Quality determinants in coffee production* (pp. 149–208). Cham: Springer International Publishing.

Chen, Z. A., Geng, Y. Y., Wang, M., Lv, D., Huang, S., Guan, Y. F., & Hu, Y. J. (2022). Relationship between microbial community and flavor profile during the fermentation of chopped red chili (*Capsicum annuum* L.). *Food Bioscience, 50*, 10. https://doi.org/10.1016/j.fbio.2022.102071

Cruz-O'Byrne, R., Piraneque-Gambasica, N., & Aguirre-Forero, S. (2021). Microbial diversity associated with spontaneous coffee bean fermentation process and specialty coffee production in northern Colombia. *International Journal of Food Microbiology, 354*, 11. https://doi.org/10.1016/j.ijfoodmicro.2021.109282

Dudareva, N., Negre, F., Nagegowda, D. A., & Orlova, I. (2006). Plant volatiles: Recent advances and future perspectives. *Critical Reviews in Plant Sciences, 25*(5), 417–440. https://doi.org/10.1080/07352680600899973

Dzialo, M. C., Park, R., Steensels, J., Lievens, B., & Verstrepen, K. J. (2017). Physiology, ecology and industrial applications of aroma formation in yeast. *FEMS Microbiology Reviews, 41*, S95–S128. https://doi.org/10.1093/femsre/fux031

El Hadi, M. A. M., Zhang, F. J., Wu, F. F., Zhou, C. H., & Tao, J. (2013). Advances in fruit aroma volatile research. *Molecules, 18*(7), 8200–8229. https://doi.org/10.3390/molecules18078200

Elhalis, H., Cox, J., & Zhao, J. (2023). Coffee fermentation: Expedition from traditional to controlled process and perspectives for industrialization. *Applied Food Research, 3*(1), 14. https://doi.org/10.1016/j.afres.2022.100253

Fan, S. S., Tang, K., Xu, Y., & Chen, S. (2020). Characterization of the potent odorants in Tibetan *Qingke Jiu* by sensory analysis, aroma extract dilution analysis, quantitative analysis and odor activity values. *Food Research International, 137*, 11. https://doi.org/10.1016/j.foodres.2020.109349

Febrianto, N. A., & Zhu, F. (2023). Coffee bean processing: Emerging methods and their effects on chemical, biological and sensory properties. *Food Chemistry, 412*, 20. https://doi.org/10.1016/j.foodchem.2023.135489

Ferreira, L. J., Gomes, M. D., de Oliveira, L. M., & Santos, L. D. (2023). Coffee fermentation process: A review. *Food Research International, 169*, 15. https://doi.org/10.1016/j.foodres.2023.112793

Fu, X. F., Li, G. P., Li, Y. A., Li, Y. Q., Bi, X. F., Huang, J. X., & Dong, W. J. (2024). Fermentation with coffee berry peels induces spatiotemporal changes in microbial communities leading to unique aroma of coffee berries. *International Journal of Food Properties, 27*(1), 657–673. https://doi.org/10.1080/10942912.2024.2339235

Góngora, C. E., Holguín-Sterling, L., Pedraza-Claros, B., Pérez-Salinas, R., Ortiz, A., & Navarro-Escalante, L. (2024). Metataxonomic identification of microorganisms during the coffee fermentation process in Colombian farms (Cesar Department). *Foods, 13*(6), 839. https://www.mdpi.com/2304-8158/13/6/839.

Happyana, N., Diniresna, A., Pratiwi, A., & Hakim, E. H. (2024). ^{1}H NMR-based metabolic profiling of green beans of *Coffea arabica* var. sigararutang with different postharvest treatments. *Journal of Food Measurement and Characterization, 11*. https://doi.org/10.1007/s11694-023-02338-0

Hou, Z. S., Wei, Y. Y., Sun, L. B., Xia, R. R., Xu, H. R., Li, Y. T., & Xin, G. (2022). Effects of drying temperature on umami taste and aroma profiles of mushrooms (*Suillus granulatus*). *Journal of Food Science, 87*(5), 1983–1998. https://doi.org/10.1111/1750-3841.16127

Huang, W. J., Fang, S. M., Wang, J., Zhuo, C., Luo, Y. H., Yu, Y. L., & Ning, J. M. (2022). Sensomics analysis of the effect of the withering method on the aroma components of Keemun black tea. *Food Chemistry, 395*, 9. https://doi.org/10.1016/j.foodchem.2022.133549

Jimenez, E. J. M., Martins, P. M. M., Vilela, A. L. D., Batista, N. N., da Rosa, S., Dias, D. R., & Schwan, R. F. (2023). Influence of anaerobic fermentation and yeast inoculation on the viability, chemical composition, and quality of coffee. *Food Bioscience, 51*, 9. https://doi.org/10.1016/j.fbio.2022.102218

Joët, T., Laffargue, A., Descroix, F., Doulbeau, S., Bertrand, B., de Kochko, A., & Dussert, S. (2010). Influence of environmental factors, wet processing and their interactions on the biochemical composition of green Arabica coffee beans. *Food Chemistry, 118*(3), 693–701. https://doi.org/10.1016/j.foodchem.2009.05.048

Kahkeshani, N., Saeidnia, S., & Abdollahi, M. (2015). Role of antioxidants and phytochemicals on acrylamide mitigation from food and reducing its toxicity. *Journal of Food Science and Technology-Mysore, 52*(6), 3169–3186. https://doi.org/10.1007/s13197-014-1558-5

Lee, K. G., & Shibamoto, T. (2002). Analysis of volatile components isolated from Hawaiian green coffee beans (*Coffea arabica* L.). *Flavour and Fragrance Journal, 17*(5), 349–351. https://doi.org/10.1002/ffj.1067

Lee, L. W., Cheong, M. W., Curran, P., Yu, B., & Liu, S. Q. (2016). Modulation of coffee aroma via the fermentation of green coffee beans with *Rhizopus oligosporus*: I. Green coffee. *Food Chemistry, 211*, 916–924. https://doi.org/10.1016/j.foodchem.2016.05.076

Liang, J. F., Wu, H. L., Lu, M. F., & Li, Y. (2024). HS-SPME-GC-MS untargeted metabolomics reveals key volatile compound changes during Liupao tea fermentation. *Food Chemistry: X, 23*, 13. https://doi.org/10.1016/j.fochx.2024.101764

Liu, C. J., Yang, Q., Linforth, R., Fisk, I. D., & Yang, N. (2019). Modifying Robusta coffee aroma by green bean chemical pre-treatment. *Food Chemistry, 272*, 251–257. https://doi.org/10.1016/j.foodchem.2018.07.226

Ma, J. Y., Li, J. P., Jin, X. L., Cesarino, I., Zeng, W., & Li, Z. (2022). Characterization of sensory properties of Yunnan coffee. *Current Research in Food Science, 5*, 1205–1215. https://doi.org/10.1016/j.crfs.2022.07.010

Marie, L., Breitler, J. C., Bamogo, P. K. A., Bordeaux, M., Lacombe, S., Rios, M., & Bertrand, B. (2024). Combined sensory, volatilome and transcriptome analyses identify a limonene terpene synthase as a major contributor to the characteristic aroma of a *Coffea arabica* L. specialty coffee. *BMC Plant Biology, 24*(1), 19. https://doi.org/10.1186/s12870-024-04890-3

Neto, D. P. D., Pereira, G. V. D., Finco, A. M. O., Letti, L. A. J., da Silva, B. J. G., Vandenberghe, L. P. S., & Soccol, C. R. (2018). Efficient coffee beans mucilage layer removal using lactic acid fermentation in a stirred-tank bioreactor: Kinetic, metabolic and sensorial studies. *Food Bioscience, 26*, 80–87. https://doi.org/10.1016/j.fbio.2018.10.005

Park, S. N., Lim, Y. K., Freire, M. O., Cho, E., Jin, D. C., & Kook, J. K. (2012). Antimicrobial effect of linalool and α-terpineol against periodontopathic and cariogenic bacteria. *Anaerobe, 18*(3), 369–372. https://doi.org/10.1016/j.anaerobe.2012.04.001

Pereira, G. V. D., Neto, D. P. D., Magalhaes, A. I., Vásquez, Z. S., Medeiros, A. B. P., Vandenberghe, L. P. S., & Soccol, C. R. (2019). Exploring the impacts of postharvest processing on the aroma formation of coffee beans - A review. *Food Chemistry, 272*, 441–452. https://doi.org/10.1016/j.foodchem.2018.08.061

Pereira, G. V. D., Neto, E., Soccol, V. T., Medeiros, A. B. P., Woiciechowski, A. L., & Soccol, C. R. (2015). Conducting starter culture-controlled fermentations of coffee beans during on-farm wet processing: Growth, metabolic analyses and sensorial effects. *Food Research International, 75*, 348–356. https://doi.org/10.1016/j.foodres.2015.06.027

Pereira, G. V. D., Sampaio, V. D., Wiele, N., Vale, A. D., Neto, D. P. D., de Souza, A. D. D., & Soccol, C. R. (2024). How yeast has transformed the coffee market by creating new flavors and aromas through modern post-harvest fermentation systems. *Trends in Food Science & Technology, 151*, 15. https://doi.org/10.1016/j.tifs.2024.104641

Pregolini, V. B., de Melo Pereira, G. V., da Silva Vale, A., de Carvalho Neto, D. P., & Soccol, C. R. (2021). Influence of environmental microbiota on the activity and

metabolism of starter cultures used in coffee beans fermentation. *Fermentation, 7*(4), 278. https://www.mdpi.com/2311-5637/7/4/278.

Procopio, S., Qian, F., & Becker, T. (2011). Function and regulation of yeast genes involved in higher alcohol and ester metabolism during beverage fermentation. *European Food Research and Technology, 233*(5), 721–729. https://doi.org/10.1007/s00217-011-1567-9

Savary, O., Coton, E., Maillard, M. B., Gaucheron, F., Le Meur, C., Frisvad, J., & Coton, M. (2023). Functional diversity of *Bisifusarium domesticum* and the newly described *Nectriaceae* cheese-associated species. *Food Research International, 168*, 13. https://doi.org/10.1016/j.foodres.2023.112691

Seninde, D. R., & Chambers, E. (2020). Coffee flavor: A review. *Beverages, 6*(3), 44. https://www.mdpi.com/2306-5710/6/3/44.

Shen, X., Yuan, W., Wang, Q., Peng, B., Guo, Y., Liu, K., & Jiang, W. (2024). Dynamic changes in microbial communities and chemical compounds during the semi-dry fermentation processing of Coffea arabica. *Fermentation, 10*(8), 435. https://www.mdpi.com/2311-5637/10/8/435.

Shofinita, D., Lestari, D., Aliwarga, L., Sumampouw, G. A., Ambarwati, S. A., Gunawan, K. C., & Achmadi, A. B. (2024). Drying methods of coffee extracts and their effects on physicochemical properties: A review. *Food and Bioprocess Technology, 17*(1), 47–72. https://doi.org/10.1007/s11947-023-03067-4

Silva, C. F., Vilela, D., Cordeiro, C. D., Duarte, W., Dias, D., & Schwan, R. (2013). Evaluation of a potential starter culture for enhance quality of coffee fermentation. *World Journal of Microbiology and Biotechnology, 29*(2), 235–247. https://doi.org/10.1007/s11274-012-1175-2

Tang, V. C. Y., Sun, J. C., Cornuz, M., Yu, B., & Lassabliere, B. (2021). Effect of solid-state fungal fermentation on the non-volatiles content and volatiles composition of *Coffea canephora* (Robusta) coffee beans. *Food Chemistry, 337*, 8. https://doi.org/10.1016/j.foodchem.2020.128023

Tat, L., Comuzzo, P., Battistutta, F., & Zironi, R. (2007). Sweet-like off-flavor in aglianico del vulture wine: Ethyl phenylacetate as the mainly involved compound. *Journal of Agricultural and Food Chemistry, 55*(13), 5205–5212. https://doi.org/10.1021/jf0637067

Thévenot, E. A., Roux, A., Xu, Y., Ezan, E., & Junot, C. (2015). Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *Journal of Proteome Research, 14*(8), 3322–3335. https://doi.org/10.1021/acs.jproteome.5b00354

Todhanakasem, T., Tai, N. V., Pornpukdeewattana, S., Charoenrat, T., Young, B. M., & Wattanachaisaereekul, S. (2024). The relationship between microbial communities in coffee fermentation and aroma with metabolite attributes of finished products. *Foods, 13*(15), 26. https://doi.org/10.3390/foods13152332

Vale, A. D. S., Pereira, C. M. T., De Dea Lindner, J., Rodrigues, L. R. S., Kadri, N. K. E., Pagnoncelli, M. G. B., & Pereira, G. V. D. M. (2024). Exploring Microbial Influence on Flavor Development during Coffee Processing in Humid Subtropical Climate through Metagenetic–Metabolomics Analysis. *Foods, 13*(12), 1871. https://www.mdpi.com/2304-8158/13/12/1871.

Vezzulli, F., Triachini, S., Mulazzi, A., Lambri, M., & Bertuzzi, T. (2022). Acrylamide: Impact of precursors concentration, origin, post-harvesting process and roasting level in high-quality arabica and Robusta coffee. *International Journal of Food Science and Technology, 57*(12), 7468–7476. https://doi.org/10.1111/ijfs.15900

Wan, L., Li, Y., Wang, H., Wang, Y. P., Song, L. P., & Liang, W. J. (2024). Rapid detection of markers in green coffee beans with different primary processing treatments of Coffea arabica L. from Yunnan. *Food Chemistry, 455*, 15. https://doi.org/10.1016/j.foodchem.2024.139942

Wan, L., Wang, H., Mo, X. Y., Wang, Y. P., Song, L. P., Liu, L. Y., & Liang, W. J. (2024). Applying HS-SPME-GC-MS combined with PTR-TOF-MS to analyze the volatile compounds in coffee husks of *Coffea arabica* with different primary treatments in Yunnan. *LWT- Food Science and Technology, 191*, 14. https://doi.org/10.1016/j.lwt.2023.115675

Wang, B. Y., Qu, F. F., Wang, P. Q., Zhao, L., Wang, Z., Han, Y. H., & Zhang, X. F. (2022). Characterization analysis of flavor compounds in green teas at different drying temperature. *LWT- Food Science and Technology, 161*, 11. https://doi.org/10.1016/j.lwt.2022.113394

Wang, C. M., Yu, J. W., Gallagher, D. L., Byrd, J., Yao, W. C., Wang, Q., & Yang, M. (2020). Pyrazines: A diverse class of earthy-musty odorants impacting drinking water quality and consumer satisfaction. *Water Research, 182*, 12. https://doi.org/10.1016/j.watres.2020.115971

Wang, J. T., Shi, J., Zhu, Y., Ma, W. J., Yan, H., Shao, C. Y., & Lin, Z. (2022). Insights into crucial odourants dominating the characteristic flavour of citrus reticulata Blanco "Chachiensis" and Camellia sinensis "Fudingdabai". *Food Chemistry, 377*, 11. https://doi.org/10.1016/j.foodchem.2022.132048

Xu, Y. Q., Wu, M. Q., Zhao, D., Zheng, J., Dai, M. Q., Li, X. T., & Sun, B. G. (2023). Simulated fermentation of strong-flavor baijiu through functional microbial combination to realize the stable synthesis of important flavor chemicals. *Foods, 12*(3), 21. https://doi.org/10.3390/foods12030644

Yuan, H. L., Cao, G. P., Hou, X. D., Huang, M. L., Du, P. M., Tan, T. T., & Luo, J. (2022). Development of a widely targeted volatilomics method for profiling volatilomes in plants. *Molecular Plant, 15*(1), 189–202. https://doi.org/10.1016/j.molp.2021.09.003

Zhai, H. N., Dong, W. J., Fu, X. F., Li, G. P., & Hu, F. G. (2024). Integration of widely targeted metabolomics and the e-tongue reveals the chemical variation and taste quality of Yunnan Arabica coffee prepared using different primary processing methods. *Food Chemistry: X, 22*, 11. https://doi.org/10.1016/j.fochx.2024.101286

Zhai, H. N., Dong, W. J., Tang, Y. M., Hu, R. S., Yu, X. X., & Chen, X. A. (2024). Characterization of the volatile flavour compounds in Yunnan Arabica coffee prepared by different primary processing methods using HS-SPME/ GC-MS and HS-GC-IMS. *LWT- Food Science and Technology, 192*, 13. https://doi.org/10.1016/j.lwt.2023.115717

Zheng, Y., Zhao, T., Wang, J., Cai, K., Chen, P., Deng, J., … Chen, M. (2022). Changes in the microbial community structure during the digitally managed fermentation of high-temperature Daqu. Food Science, 43*(12)*, 171–178. <Go to ISI>://CSCD: 7235811.

Zhou, J. P., Chen, L. F., Foo, H. L., Cao, Z. H., & Lin, Q. Y. (2024). Changes in microbial diversity and volatile metabolites during the fermentation of Bulang pickled tea. *Food Chemistry, 458*, 16. https://doi.org/10.1016/j.foodchem.2024.140293

Zhu, Y., Lv, H. P., Shao, C. Y., Kang, S. Y., Zhang, Y., Guo, L., & Lin, Z. (2018). Identification of key odorants responsible for chestnut-like aroma quality of green teas. *Food Research International, 108*, 74–82. https://doi.org/10.1016/j.foodres.2018.03.026